HapFGAPI: Multimodal Framework for Gesture and Facial Expression Analysis

Christine Lisetti
SCIS
Florida International University
Miami, Florida, USA
lisetti@cs.fiu.edu

Wei Zeng
SCIS
Florida International University
Miami, Florida, USA
wzeng@cs.fiu.edu

Md. Shakil Ahamed
SCIS
Florida International University
Miami, Florida, USA
maham001@cs.fiu.edu

Abstract

HapFGAPI is a framework for developing Facial expressions, Gesture animations and Text to speech using Virtual Character (VC). This is an integrated platform consists of command based VC , 3D Environment. Multi-modal signals are taken from input devices. The input options are AcceleGlove and Kinetic 3D camera for taking multi-modal signals. The main intention for this framework is to generate platform for American Sign Language(ASL) research.

1 Introduction

For various research applications virtual human character are getting popular. For example electronic commerce [], virtual health coach[], for various entertainment applications like video games[]. Intelligent decision making, emotion, proactive and reactive behaviors makes the VC oriented applications more acceptable instead of real human[].

Sign language generation, translation, communication is a large research domain. Virtual characters are being used for this application area. It is used extensively in education and communication easiness for a group of people.

Animation are not only confined into moving objects or games or movie development only but used for showing affective nature of human or moving objects[]. Animations is used for generating user's signs. Gestures and facial expressions are parts of sign. Some tools are used to generate those sign animations[]. But animation quality, realistic human model, facial expression in the model are not adequate. At the same time their is no 3D environment where VC can move and control by means of commands.

From this motivation, Virtual character animation framework (Application Programming Interface) and 3D platform becomes visibly crucial for research community in perspective to the native and other signers for successful social exchange of communication.

This research focus on a complete platform development by creating an integrated VC and 3D environment framework whether researcher can focus more on applications and their usages analysis rather consuming time for making development settings.

1.1 Literature Review

The research focus of this framework consists of two domains. First domain is about graphics. Various research are done on 3D avatar's body gesture and Sensing human body signals. And Second domain is about American Sign Language(ASL). In this section, authors focus on the state of art of 3D virtual characters on ASL research.

In 2013, Microsoft Research, Yupeng Zhang, Teng Han, et.all published "Body Avatar" paper. It is a Kinect-based system to create 3D avatar using human body gesture which allow users without professional expertises.They used two metaphors, first and third person. First person is the dominant user and third is for generating those shapes which is less convenient for first user. They used tree shaped skeleton for the VC. Triangular mesh used for meta ball model representation which is sophisticated for free motion generations. They used different editing operations for generating attractive avatars.

In 2011, University of Pennsylvania, Joseph T. Kider Jr., Kaitlin Pollock, and Alla Safonova, published their paper on "Data driven appearance model for human fatigue". They showed how to visualize humans breathe, flush and sweat when they perform physical activities. A Bio-signal capture technique have been used Five sensors for taking input data of respiration,cardiograph, skin temperature, blood

pulse, skin. They captured the data, cleaned the motion capture take, applied the biodata to the simulation parameters, deformed the organs, which in turn deformed the skin, change the appearance of the figure (sweating, flushing), then render the final output. Yet some limitations their model visualized fatigued person clearly.

In 2013, University of Southern California; Alexander, Oleg Fyffe, Graham Busch, Jay Yu, published "Digital Ira: Creating a Real-Time Photoreal Digital Actor" where they created a realistic virtual character. They used surface space value based real time rendering and scans are stored as images. Hence the appearance of the model could contain realistic wrinkles and advanced lighting effect.

In 2011, City University of New York, Matt Huenerfauth, Pengfei Lu, Andrew Rosenberg, published paper [1] where they made empirical evaluation of the face movement impact on understandability and perceived quality of ASL or PSE (Pidgin Sign English). The contribution of this research was to understand the impact of animated character's facial expression on ASL and PSE sign users Analysis of facial expression for generating grammar. They performed their evaluation on native signers. They used vcom3D animation api. But these tool is not realistic and they used image as their background.

1.2 Tools and Library used

To implement the framework some well known tools have been used. Name and description of each of the tools are given below: In our system authors use Haptek API which is available in market (www.haptek.com) which contains hypertext, tts, commands facilities. Authors also use Blender for 3D modeling, webGL for making the content applicable on html page. To make the sysetm dynamic they used Javascript and .NET language.

Logo	Tools	Description
	Haptek API	API-tools, hypertext, tts, commands
	Blender	Open Source, 3D modeler
	WebGL API	WebGL (Web Graphics Library) is a JavaScript API for rendering interactive 3D graphics and 2D graphics
	HTML5	Latest HTML markup language
	JavaScript	JavaScript is the scripting language of the Web.

Table 1: A table caption.

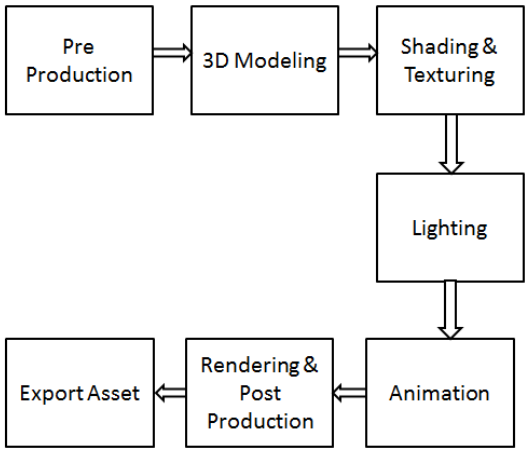

Figure 1: 3D Background *Modeling* Pipeline.

2 Modeling Solution

Modeling the framework three steps need to complete.

2.1 3D Background Modeling

To model the background the following steps should be completed.

2.2 3D Character Modeling

To model 3d Virtual Character the following steps should be completed.

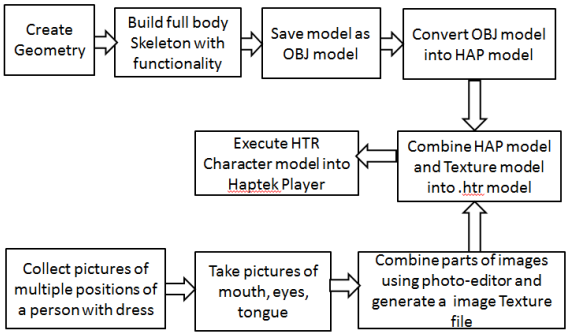

Figure 2: 3D Character *Modeling* Pipeline.

2.3 Animation Generation

To generate animations the following steps should be completed.

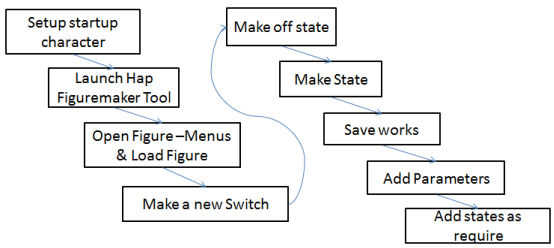

Figure 3: Animations Pipeline.

3 Implementation Details

3.1 Integration of Virtual Character in 3D Room

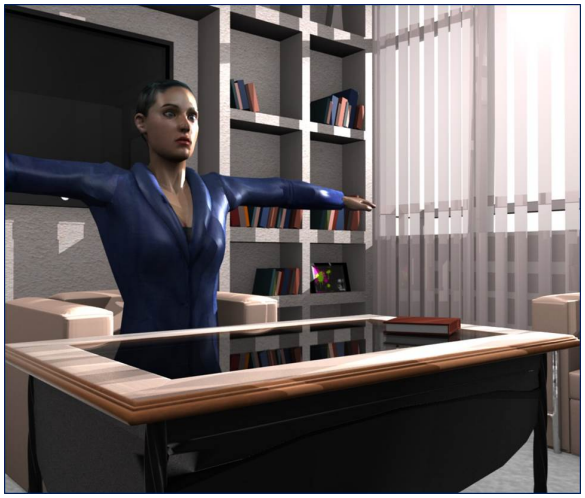

Figure 4: Sample 3D Room Design

3.2 Right Hand Shapes for sign generation

Figure 5: Sample 3D Room Design

3.3 Words ASL Animations

Figure 6: 10 Words animations

3.4 Project Implementation Schedule

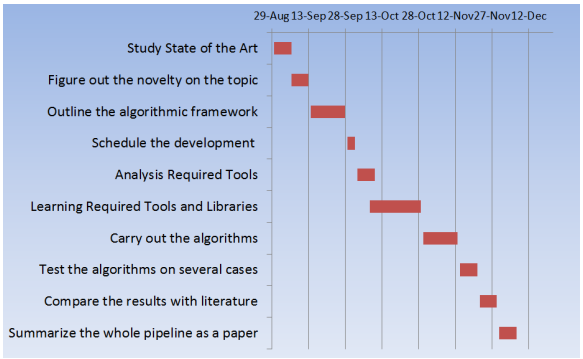

Figure 7: Project Schedule

3.5 Next Goal

Complete ASL Gesture Code Generation Interfacing with Input scan signal and codes

3.6 Long Term Goal

Create a computational model of social multimodal interactive dialog signals. Adapting to the users multimodal signals in realtime Character Breathing. Character Flushing. Port/adapt Haptek characters situated in 3D scenes in webGL Independent Gesture Control Framework.

4 Summary and Conclusions

This template will get you through the minimum article, i.e. no figures or equations. To include those, please refer to your LaTeX manual and the IEEE publications guidelines. Good Luck!

Acknowledgements

This is how to do an unnumbered subsection, which comes out in 11 point bold font. Here I thank my colleagues, especially Mike Gennert, who know more about Tex and LaTeX than I.

References

[1] Matt Huenerfauth, Pengfei Lu, Andrew Rosenberg, "Evaluating Importance of Facial Expression in American Sign Language and Pidgin Signed English Animations Authors," *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*,pp. 99-106, ACM New York, NY, USA,1987.

[2] Santiago Martinez, Robin J.S. Sloan, Andrea Szymkowiak and Ken Scott-Brown, "Animated Virtual Agents to Cue User Attention," *CORE, Journal Article, published, peer-reviewed*, 2011.

[3] Yupeng Zhang, Teng Han, Zhimin Ren, N Umetani, X Tong, Y Liu, T Shiratori, X Cao," BodyAvatar: creating freeform 3D avatars using first-person body gestures," *In Proceedings of the 26th annual ACM symposium on User interface software and technology,* pp. 387-396, 2013.

[4] Joseph T. Kider Jr., Kaitlin Pollock, and Alla Safonova, "A Data-driven Appearance Model for Human Fatigue," *n Proceedings of the Eurographics/ ACM SIGGRAPH Symposium on Computer Animation* , 2011.

[5] Alexander, Oleg Fyffe, Graham Busch, Jay Yu, "Digital Ira: Creating a Real-Time Photoreal Digital Actor," *In Proceedings of the ACM SIGGRAPH*, 2013.

[6] Ryosuke Ichikari, Oleg Alexander, Paul Debevec, "Vuvuzela: A Facial Scan Correspondence Tool,"

In Proceedings of the ACM SIGGRAPH,Posters, 2013.

[7] Reza Amini, C. Lisetti, "HapFACS: an Open Source API/Software to Generate FACS-Based Expressions for ECAs Animation andfor Corpus Generation," *5th biannual Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII 2013)*, IEEE Computer Society, 2013.